<u>WHAT IS CLAIMED IS</u>:

1.   A method for analyzing data from hybridization of a sample to an array of oligonucleotide probes,

wherein the sample comprises a plurality of nucleotide sequences, each nucleotide sequence corresponding to a particular gene,

wherein some or all of the oligonucleotide probes are assigned to invertible subblocks such that each gene which hybridizes to an oligonucleotide probe assigned to a particular subblock does not hybridize to the oligonucleotide probes in the other subblocks, and

which method comprises a step of separately analyzing the data for the oligonucleotide probes in each subblock.

2.   A method according to claim 1, wherein the array comprises a plurality (N_0) of oligonucleotide probes having a particular sequence length n so that all nucleic acid sequences having the particular sequence length are present on the array.

3.   A method according to claim 2 wherein the particular sequence length n is from about 6 to about 20.

4.   A method according to claim 3 wherein the particular sequence length n is from about 9 to about 16.

5.   A method according to claim 4 wherein the particular sequence length n is from about 10 to about 12.

6.   A method according to claim 4 wherein the particular sequence length n is from about 12 to about 15.

7.   A method according to claim 1 wherein oligonucleotide probes are assigned to subblocks according to a method which comprises, for each subblock, steps of:

(a)   associating a gene g_a with a gene list for a subblock, wherein the gene g_a is not already associated with a gene list for a subblock; and

(b)   assigning an oligonucleotide probe o_x to the subblock, wherein the oligonucleotide probe o_x hybridizes to the gene g_a,

wherein the steps are repeated for each subblock until each gene is associated with a gene list for a subblock.

8.   A method according to claim 7, further comprising steps of:

(c)   for each probe o_x assigned to the subblock, associating genes g_b with the gene list for the subblock, wherein each gene g_b hybridizes to the probe o_x; and

(d)   for each gene g_b associated with the gene list, assigning an oligonucleotide probe o_y to the subblock, wherein the oligonucleotide probe o_y hybridizes to the gene g_b.

9.   A method according to claim 8 wherein the steps of:

(c)   associating genes g_b with the gene list for the subblock; and

(d)   assigning an oligonucleotide probe o_y for each gene g_b associated with the gene list

are iteratively repeated for each oligonucleotide probe o_y assigned in step (d).

10.   A method according to claim 9 wherein the steps (c) - (d) are repeated for not more than 100 iterations.

11.   A method according to claim 10 wherein the steps (c) - (d) are repeated for not more than 50 iterations.

12.   A method according to claim 11 wherein the steps (c) - (d) are repeated for not more than 20 iterations.

13.   A method according to claim 12 wherein the steps (c) - (d) are repeated for not more than 15 iterations.

14.   A method according to claim 13 wherein the steps (c) - (d) are repeated for not more than ten iterations.

15.   A method according to claim 14 wherein the steps (c) - (d) are repeated for not more than five iterations.

16.   A method according to claim 15 wherein the steps (c) - (d) are repeated for not more than four iterations.

17.   A method according to claim 16 wherein the steps (c) - (d) are repeated for not more than three iterations.

18.   A method according to claim 17 wherein the steps (c) - (d) are repeated for not more than two iterations.

19.   A method according to claim 9 wherein the steps (c) - (d) are iteratively repeated until, for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

20.   A method according to claim 8 wherein:

    (i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe; and

    (ii)   the steps (c) - (d) are iteratively repeated until for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

21.   A method according to claim 7 in which:

    (i)   each oligonucleotide probe assigned to the subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, the degeneracy value being equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock; and

    (ii)   each gene g_a associated with the gene list for the subblock hybridizes to at least one oligonucleotide probe o_x having a degeneracy value less than the particular threshold T.

22.   A method according to claim 1 wherein:

    (i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, and

    (ii)   the degeneracy value is equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock.

23.   A method according to claim 22 wherein the particular threshold T is no more than 100.

24.   A method according to claim 23 wherein the particular threshold T is no more than 50.

25.   A method according to claim 24 wherein the particular threshold T is no more than 20.

26.   A method according to claim 25 wherein the particular threshold T is no more than ten.

27.   A method according to claim 26 wherein the particular threshold T is no more than five.

28.   A method according to claim 27 wherein the particular threshold T is no more than four.

29.   A method according to claim 28 wherein the particular threshold T is no more than three.

30.   A method according to claim 29 wherein the particular threshold T is no more than two.

31.   A method according to claim 30 wherein the particular threshold T is one.

32.   A method according to claim 1 in which expression levels are determined for each gene g_i that hybridizes to oligonucleotide probes assigned to a particular subblock by a method which comprises solving a system of linear equations for the hybridization of each gene g_i to each oligonucleotide probe o_j assigned to the particular subblock.

33.   A method according to claim 32 wherein the system of linear equations is of the form

$$\vec{E} = (\hat{H}')^{-1} \cdot \vec{S}'$$

wherein:

(a)   each element E_i of the vector $\vec{E}$ indicates abundance of a nucleotide sequence in the sample corresponding to a particular gene g_i;

(b)   each element S_j of the vector $\vec{S}'$ indicates a level of hybridization to a particular oligonucleotide probe o_j; and

(c)   each element H_{ij} of the matrix $\hat{H}'$ indicates hybridization affinity of the nucleotide sequence corresponding to said particular particular gene g_i for the particular oligonucleotide probe o_j.

34.   A method according to claim 1 wherein each of the nucleotide sequences has a length ℓ_i equal to the length of the corresponding gene.

35.   A method according to claim 1 wherein the length of each different nucleic acid is decreased before hybridization so that each different nucleic acid has a decreased length $L_i = \ell_i - \Delta L_i$ that is less than the length of the corresponding gene.

36.   A method according to claim 35 wherein the length is decreased by enzymatic digestion.

37.   A method according to claim 35 wherein the length of each different nucleic acid is decreased, on average, by a controled amount $<\Delta L>$.

38.   A method according to claim 37 wherein the amount $<\Delta L>$ is between about 50 and about 500 bases.

39.   A method according to claim 38 wherein the amount $<\Delta L>$ is between about 50-100 bases.

40.   A method according to claim 38 wherein the amount $<\Delta L>$ is between about 100-500 bases.

41.   A method according to claim 35 wherein the length of each different nucleic acid is decreased by a method which comprises:

      (i)   protecting each nucleic acid along a particular length; and

      (ii)   removing the unprotected portion.

42.   A method according to claim 35 wherein the average decreased length $<L>$ is controled.

43.   A method according to claim 42 wherein the average decreased length $<L>$ is less than or equal to about 500 bases.

44.   A method according to claim 43 wherein the average decreased length $<L>$ is less than or equal to about 100 bases.

45.   A method according to claim 44 wherein the average decreased length $<L>$ is about 50 bases.

46.   A method according to claim 42 wherein the average decreased length $<L>$ is between about 50 and 100 bases.

47.   A method according to claim 42 wherein the average decreased length $<L>$ is between about 100 and 500 bases.

48. A method for assigning all or some of a plurality of oligonucleotide probes to subblocks suitable for analyzing data from hybridization of a sample to an array of the oligonucleotide probes,

wherein the sample comprises a plurality of nucleotide sequences, each nucleotide sequence corresponding to a particular gene,

which method comprises steps of:

(a) associating a gene g_a with a gene list for a subblock, wherein the gene g_a is not already associated with a gene list for a subblock; and

(b) assigning an oligonucleotide probe o_x to the subblock, wherein the oligonucleotide probe o_x hybridizes to the gene g_a,

wherein the steps are repeated for each subblock until each gene is associated with a gene list for a subblock.

49. A method according to claim 48 further comprising steps of:

(c) for each probe o_x assigned to the subblock, associating genes g_b with the gene list for the subblock, wherein each gene g_b hybridizes to the probe o_x; and

(d) for each gene g_b associated with the gene list, assigning an oligonucleotide probe o_y to the subblock, wherein the oligonucleotide probe o_y hybridizes to the gene g_b.

50. A method according to claim 49 wherein the steps of:

(c) associating genes gb with the gene list for the subblock; and

(d) assigning an oligonucleotide probe o_y for each gene g_b associated with the gene list

are iteratively repeated.

51.   A method according to claim 50 wherein the step (c) - (d) are repeated for not more than 100 iterations.

52.   A method according to claim 51 wherein the steps (c) - (d) are repeated for not more than 50 iterations.

53.   A method according to claim 52 wherein the steps (c) - (d) are repeated for not more than 20 iterations.

54.   A method according to claim 53 wherein the steps (c) - (d) are repeated for not more than ten iterations.

55.   A method according to claim 54 wherein the steps (c) - (d) are repeated for not more than five iterations.

56.   A method according to claim 55 wherein the steps (c) - (d) are repeated for not more than four iterations.

57.   A method according to claim 56 wherein the steps (c) - (d) are repeated for not more than three iterations.

58.   A method according to claim 57 wherein the steps (c) - (d) are repeated for not more than two iterations.

59.   A method according to claim 50 wherein the steps (c) - (d) are iteratively repeated until, for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

60.   A method according to claim 51 wherein:

(i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe; and

(ii)   the steps (c) - (d) are iteratively repeated until, for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

61.   A method according to claim 48 in which:

(i)   each oligonucleotide rpboe assigned to the subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, the degeneracy value being equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock; and

(ii)   each gene g_a associated with the gene list for the subblock hybridizes to at least one oligonucleotide probe o_x having a degeneracy value less than the particular threshold T.

62.   A method according to claim 48 wherein:

(i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, and

(ii)   the degeneracy value is equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock.

63.   A method according to claim 48 , wherein the array comprises a plurality (N_0) of oligonucleotide probes having a particular sequence length n so that all nucleic acid sequences having the particular sequence length are present on the array.

-58-

64.   A method for selecting a particular sequence length n for an array comprising a plurality (N_0) of oligonucleotide probes having the particular sequence length n, which method comprises:

(a)   identifying a sequence length n providing an average probe degeneracy $<d(n)>$ suitable for analyzing nucleic acid expression using the array; and

(b)   selecting the identified sequence length n,

wherein the average probe degeneracy $<d(n)>$ indicates the number of different nucleic acids that hybridize, on average, to a particular oligonucleotide probe.

65.   A method according to claim 64 wherein each of the different nucleic acids corresponds to a gene in a plurality(N_g) of different genes.

66.   A method according to claim 65 wherein each of the nucleotide sequences has a length ℓ_i equal to the length of the corresponding gene.

67.   A method according to claim 66 wherein the length of each different nucleic acid is decreased before hybridization so that each different nucleic acid has a decreased length $L_i = \ell_i - \Delta L_i$ that is less than the length of the corresponding gene.

68.   A method according to claim 67 wherein the length is decreased by enzymatic digestion.

69.   A method according to claim 67 wherein the length of each different nucleic acid is decreased, on average, by a controled amount $<\Delta L>$.

70.   A method according to claim 69 wherein the amount $<\Delta L>$ is between about 50 and about 500 bases.

71.   A method according to claim 70 wherein the amount $<\Delta L>$ is between about 50-100 bases.

72.   A method according to claim 70 wherein the amount $<\Delta L>$ is between about 100-500 bases.

73.   A method according to claim 67 wherein the length of each different nucleic acid is decreased by a method which comprises:

      (i)   protecting each nucleic acid along a particular length; and

      (ii)   removing the unprotected portion.

74.   A method according to claim 67 wherein the average decreased length $<L>$ is controled.

75.   A method according to claim 74 wherein the average decreased length $<L>$ is less than or equal to about 500 bases.

76.   A method according to claim 75 wherein the average decreased length $<L>$ is less than or equal to about 100 bases.

77.   A method according to claim 76 wherein the average decreased length $<L>$ is about 50 bases.

78.   A method according to claim 74 wherein the average decreased length $<L>$ is between about 50 and 100 bases.

79.   A method according to claim 74 wherein the average decreased length $<L>$ is between about 100 and 500 bases.

80.   A method according to claim 64 wherein the nucleic acids hybridize to the oligonucleotide probes with no more than a particular number (m) of base-pair mismatches.

81.   A method according to claim 80 wherein the average probe degeneracy $<d(n)>$ is provided by the relation

$$\langle d(n) \rangle = \frac{N_g}{N_0}(1 - n + \langle L \rangle) \times c$$

c is provided by the relation

$$c = \sum_{k=0}^{m} \binom{n}{k} 3^k \text{, and}$$

$<L>$ indicates the average length of the different nucleic acids.

82.   A method according to claim 64 wherein the average probe degeneracy $<d(n)>$ is provided by the relation

$$\langle d(n) \rangle = \frac{N_g}{N_0}(1 - n + \langle L \rangle) \times c$$

wherein $<L>$ indicates the average length of the different nucleic acids, and c indicates the number of the different nucleic acids that hybridize, on average, to an oligonucleotide probe having the particular sequence length n.

83.   A method according to claim 64 wherein the step (a) of identifying a sequence length n comprises:

   (i)   comparing oligonucleotide sequences having a particular sequence length n with sequences of the different nucleic acids, so that

nucleic acids which hybridize to each oligonucleotide sequence are identified; and

(ii)   determining the average probe degeneracy $\langle d(n)\rangle$ from the number of different nucleic acids that hybridize to each oligonucleotide sequence.

84.   A method according to claim 64 wherein the identified sequence length n provides an average probe degeneracy $\langle d(n)\rangle$ that is less than or equal to about five.

85.   A method according to claim 64 wherein the identified sequence length n provides an average probe degeneracy $\langle d(n)\rangle$ that is less than or equal to about four.

86.   A method according to claim 64 wherein the identified sequence length n provides an average probe degeneracy $\langle d(n)\rangle$ that is less than or equal to about three.

87.   A method according to claim 64 wherein the identified sequence length n provides an average probe degeneracy $\langle d(n)\rangle$ that is less than or equal to about two.

88.   A method according to claim 64 wherein the identified sequence length n provides an average probe degeneracy $\langle d(n)\rangle$ of about one.

89.   A method according to claim 64 wherein the step (a) of identifying a sequence length n comprises:

(i)   assigning all or some of a plurality of oligonucleotide probes having a particular sequence length n to subblocks suitable for analyzing data from hybridization of a sample to an array of the oligonucleotide probes; and

(ii)   determining the average probe degeneracy $\langle d(n)\rangle$ from the oligonucleotide probes assigned to the subblocks.

90.   A method according to claim 89 wherein the plurality of oligonucleotide probes is a plurality all nucleic acid sequences having the particular length n.

91.   A method according to claim 89 wherein the oligonucleotide probes are assigned to subblocks according to a method which comprises steps of:

(a)   associating a gene g_a with a gene list for a subblock, wherein the gene g_a is not already associated with a gene list for a subblock; and

(b)   assigning an oligonucleotide probe o_x to the subblock, wherein the oligonucleotide probe o_x hybridizes to the gene g_a,

wherein the steps are repeated for each subblock until each gene is associated with a gene list for a subblock.

92.   A method according to claim 91 further comprising steps of:

(c)   for each probe o_x assigned to the subblock, associating genes g_b with the gene list for the subblock, wherein each gene g_b hybridizes to the probe o_x; and

(d)   for each gene g_b associated with the gene list, assigning an oligonucleotide probe o_y to the subblock, wherein the oligonucleotide probe o_y hybridizes to the gene g_b.

93.   A method according to claim 92 wherein the steps of:

(c)   associating genes gb with the gene list for the subblock; and

(d)   assigning an oligonucleotide probe o_y for each gene g_b associated with the gene list

are iteratively repeated.

94.   A method according to claim 93 wherein the step (c) - (d) are repeated for not more than 100 iterations.

-63-

95.   A method according to claim 94 wherein the steps (c) - (d) are repeated for not more than 50 iterations.

96.   A method according to claim 95 wherein the steps (c) - (d) are repeated for not more than 20 iterations.

97.   A method according to claim 96 wherein the steps (c) - (d) are repeated for not more than ten iterations.

98.   A method according to claim 97 wherein the steps (c) - (d) are repeated for not more than five iterations.

99.   A method according to claim 98 wherein the steps (c) - (d) are repeated for not more than four iterations.

100.   A method according to claim 99 wherein the steps (c) - (d) are repeated for not more than three iterations.

101.   A method according to claim 101 wherein the steps (c) - (d) are repeated for not more than two iterations.

102.   A method according to claim 93 wherein the steps (c) - (d) are iteratively repeated until, for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

103.   A method according to claim 93 wherein

(i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe; and

(ii)   the steps (c) - (d) are iteratively repeated until, for each oligonucleotide probe o_x assigned to the particular subblock, all genes g_a that hybridize to the oligonucleotide probe o_x are associated with the gene list for the particular subblock.

104.   A method according to claim 91 in which:

(i)   each oligonucleotide probe assigned to the subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, the degeneracy value being equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock; and

(ii)   each gene g_a associated with the gene list for the subblock hybridizes to at least one oligonucleotide probe o_x having a degeneracy less than the particular threshold T.

105.   A method according to claim 91 wherein:

(i)   each oligonucleotide probe assigned to a subblock has a degeneracy value indicating the number of different genes that hybridize to that oligonucleotide probe, and

(ii)   the degeneracy value is equal to or below a particular threshold T for each oligonucleotide probe assigned to the subblock.